

ПРОБЛЕМИ АВТОМАТИЗАЦІЇ НАДАННЯ РЕКОМЕНДАЦІЙ

Гольштейн М.М., Бабкова Н.В., Угольнікова Н.С.

*Національний технічний університет
«Харківський політехнічний інститут»,
м. Харків*

Поява рекомендаційних систем на сучасному ІТ ринку обумовлена заміною статичного списку рекомендацій при пошуку або покупках на веб-сайтах. Рекомендаційною системою називають підклас систем фільтрації інформації, що будують рейтингові переліки об'єктів (фільмів, музики, книг, новин, веб-сайтів), яким користувач може надати перевагу. Для цього використовується інформація з профілю користувача.

В наш час будь-які рекомендації на основі коментарів користувачів є практичною задачею. Раніше всі відомості щодо предмету рекомендації зберігались на паперах, ще було ризиком через їх велику кількість та можливість втрати. Зі стрімким розвитком електронних програм та інформаційних систем ця задача значно полегшилася. Коментарі для рекомендації стали доступними на форумах, в соціальних мережах, в YouTube та інших додатках. Сучасні рекомендаційні системи мають ряд стандартних проблем та недоліків, дослідження яких та розробка методів їх подолання є актуальною науково-практичною задачею:

1) проблема холодного старту (Cold-start Problem, CSP) виникає тоді, коли в системі з'являються нові елементи або нові користувачі (User Cold-Start), історія вподобань яких порожня, або нові об'єкти (Item Cold-Start), у яких ще немає оцінок та/або набору ознак;

2) проблема бульбашки фільтрів. Класичні рекомендаційні системи пропонують користувачам об'єкти, виходячи лише з їх попередніх вподобань. Отже, користувач потрапляє у інформаційне середовище, в якому спостерігає лише обмежену кількість однотипних об'єктів;

3) проблема розріджених даних зазвичай виникає при оцінці користувачами обмеженої кількості доступних елементів, особливо коли каталог великий. Результатом є розріджена матриця рейтингу користувачів з недостатніми даними для ідентифікації подібних користувачів або товару, що негативно впливає на якість рекомендацій;

4) шахрайство у рекомендаційних системах, де кожен може ставити оцінки, люди можуть давати позитивні оцінки своїм предметам і погані своїм конкурентам;

5) проблема синонімії полягає в тенденції схожих і однакових предметів мати різні імена, наприклад, «музика для дітей» та «дитяча музика» відносяться до одного жанру, але система сприймає їх як різні.

6) проблема «білих ворон», до яких відносяться користувачі, чия думка постійно не збігається з більшістю інших. Через унікальність смаку їм неможливо щось рекомендувати.

Іншими проблемами є відсутність персоналізації, збереження конфіденційності, зменшення шуму, інтеграція джерел даних, відсутність новизни та адаптивність переваг користувача.